

INFORMES DE INVESTIGACIÓN

IDALITH LEON ORTIZ, M.*

EL GESE: UN GENERADOR DE ENUNCIADOS SIMPLES EN ESPAÑOL BASADO EN EL MODELO SEMÁNTICO PRAGMÁTICO DE LA LINGÜÍSTICA APLICADA¹

Resumen

*La investigación que aquí se presenta tiene como resultado un dispositivo informático: el GESE **Generador de Enunciados Simples en Español**. Dentro del ámbito de la Lingüística Computacional, se denomina "generador" a un programa de computador capaz de producir enunciados en "lenguaje natural" a partir de inferencias. Para lograr tal fin, el generador desarrollado tiene como sustento lingüístico el llamado Modelo Semántico - Pragmático, perteneciente a la lingüística aplicada, que a su vez se fundamenta en la Semántica Generativa. El fuerte de la investigación radica en el desarrollo tecnológico mismo, así como en la validación del modelo Semántico Pragmático en un nuevo campo del saber: la Lingüística Computacional. La investigación demuestra que dicho modelo es adecuado y potente como soporte para desarrollar programas orientados al procesamiento computacional de lenguaje natural, más específicamente en el área de estudio llamada "generación de lenguaje" (language generation). Por otra parte, la aplicación tecnológica presentada (GESE) permite verificar que PROLOG es una excelente herramienta para el desarrollo de este tipo de programas.*

Presentación

Dentro de la gran variedad de áreas que cubren los estudios de Inteligencia Artificial (IA), adquiere cada vez más relevancia la llamada Lingüística Computacional (LC) (Hays: Computational Linguistics), es decir, el campo de investigación dedicado al estudio y tratamiento computarizado del Lenguaje Natural².

* Licenciada en Español-Inglés de la Universidad Pedagógica Nacional: Maestría en Lingüística del Instituto Caro y Cuervo; Maestría en Tecnologías de la Información Aplicadas a la Educación. Profesora de Ciencia Cognitiva de la Especialización en Informática para Docentes de la Universidad Manuela Beltrán.

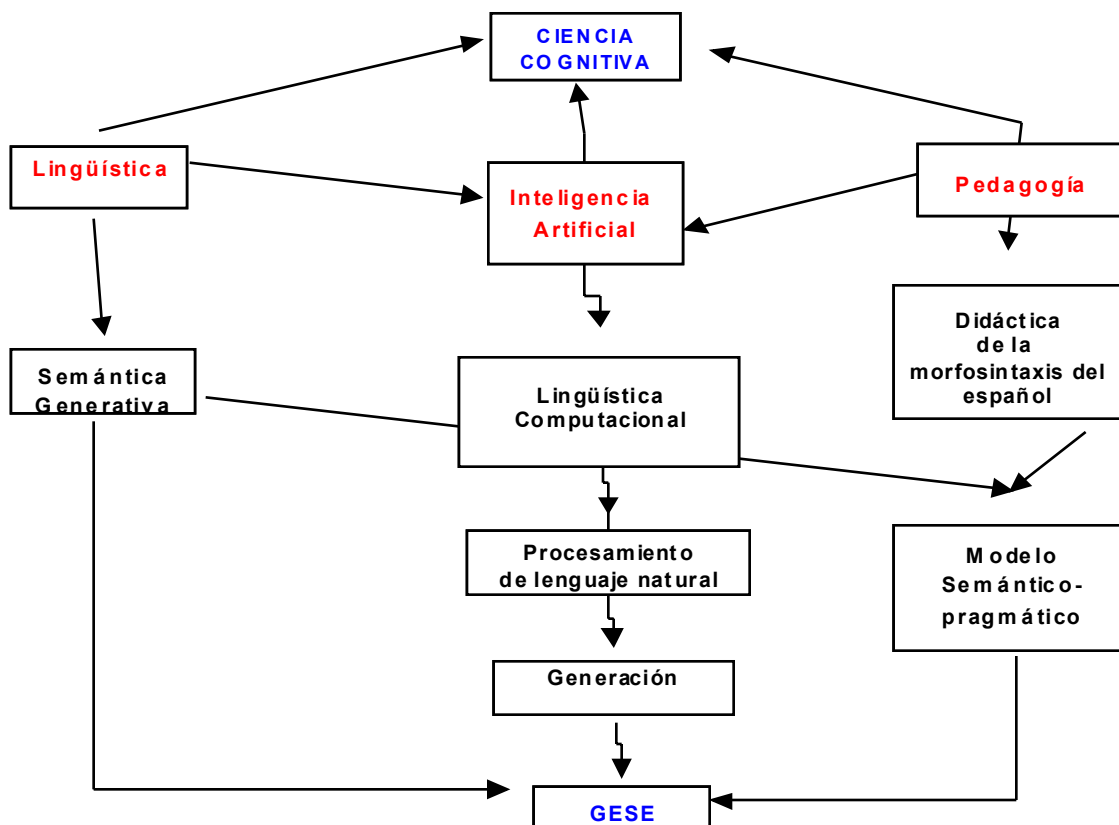
¹ Investigación desarrollada en el marco de la maestría en Tecnologías de la información aplicadas a la educación, ofrecida por el Departamento de Tecnología de la Universidad Pedagógica Nacional.

² Es necesario aclarar que dentro de la Teoría Lingüística se reserva la denominación de **Lenguaje** para referirse a la facultad del ser humano, que le permite abstraer la realidad, simbolizarla y, si desea, expresarla.

El origen de esta rama de la informática es tan antiguo como el de los ordenadores; en efecto los primeros trabajos se remontan a los años cuarenta, cuando fue posible tratar textos escritos a través del ordenador, gracias a que se había extendido su capacidad para manipular símbolos. Se pueden considerar como pioneros los trabajos de traducción automática de A.D. Booth y W. Weaver (1940-1949).

Aunque los trabajos iniciales se limitaron a crear índices, listas de palabras, concordancias, los trabajos posteriores y más aún los que se adelantan actualmente, alcanzan complejos y variados campos de aplicación, que van desde los bien conocidos procesadores de palabras (Wordperfect, Wordstar, Words) hasta las interfaces hombre-máquina, pasando por edición automatizada de diccionarios, sistemas de documentación automática, narraciones de historias, modelos de adquisición de la lengua, procesamiento del habla, etc.

Es obvia la función que ha cumplido la Lingüística Teórica (LT) dentro del desarrollo de la LC, pero así mismo debe reconocerse que gracias a la última han surgido métodos y modelos que han revertido en teorías y modelos de la LT; además la LC se ha constituido en excelente herramienta de validación de las teorías y modelos lingüísticos. Se debe hablar, entonces, de una relación biunívoca entre Lingüística Teórica y la Lingüística Computacional. Y es en la intersección de estas áreas del conocimiento que se sitúa el trabajo de investigación que nos ocupa, en el que confluye, además, la pedagogía, en la medida en que incorpora elementos relacionados con la didáctica de la morfosintaxis del español. Así, puede definirse como una investigación de carácter formal, enmarcada en la Ciencia Cognitiva, área del conocimiento que engloba los saberes mencionados, como se representa la figura 1.



Segunda época, No. 11 – Segundo Semestre de 1999

Fig. 1. Marco de la investigación

Bajo esta perspectiva, se plantearon los siguientes objetivos de investigación:

Generales

- Contribuir al avance de las investigaciones tanto en el campo de la lingüística computacional como de la lingüística aplicada.
- Verificar si el Semántico - Pragmático, como modelo de análisis lingüístico, es adecuado para la creación de modelos computacionales en el área del procesamiento del lenguaje natural, específicamente en el campo de la generación.

Específicos

- Desarrollar un generador de enunciados simples en español que, además de presentar el contenido léxico de los enunciados, represente la estructura de sus componentes semántico, sintáctico y pragmático y así simule el proceso de generación de enunciados simples por parte de un hablante real de esta lengua.
- Diseñar un modelo computacional para el desarrollo de un generador de enunciados simples en español, a partir del modelo Semántico – pragmático de la Lingüística aplicada.

El proceso investigativo

La investigación en el campo de la Lingüística Computacional, en general, y de la Generación en Lenguaje Natural, en particular, puede adquirir diversos matices, dentro de los cuales se encuentra tanto el desarrollo de investigaciones formales, como el desarrollo de herramientas tecnológicas. Dado que es el caso de la investigación realizada, a continuación se describen los componentes que intervinieron para su desarrollo.

Componentes del proceso investigativo.

1. Identificación y descripción de la perspectiva teórica que fundamenta la propuesta.
2. Selección y descripción exhaustiva del modelo lingüístico base.
3. Selección de lenguaje de programación.
4. Diseño del modelo computacional
5. Desarrollo del generador.
6. Validación
7. Conclusiones
8. Proyecciones

Aunque cada uno de estos componentes tiene su propia especificidad, entre ellos debe existir una estrecha relación que garantice la coherencia del proceso global. Es así como el modelo lingüístico seleccionado como base para el desarrollo del modelo computacional, y por lo tanto del generador, debe responder a la teoría lingüística que soporta toda la propuesta. Este hecho muestra cómo los dos primeros componentes del proceso entran en escena de manera simultánea.

En cuanto al tercer componente, la selección del lenguaje de programación, en el caso que nos ocupa se tenía previsto desde el inicio de la investigación, sin embargo, puede darse en otro momento del proceso, dependiendo de su mayor o

menor grado de incidencia en otro(s) componente(s). Eso sí, debe estar definido para poder iniciar el desarrollo tecnológico propiamente dicho.

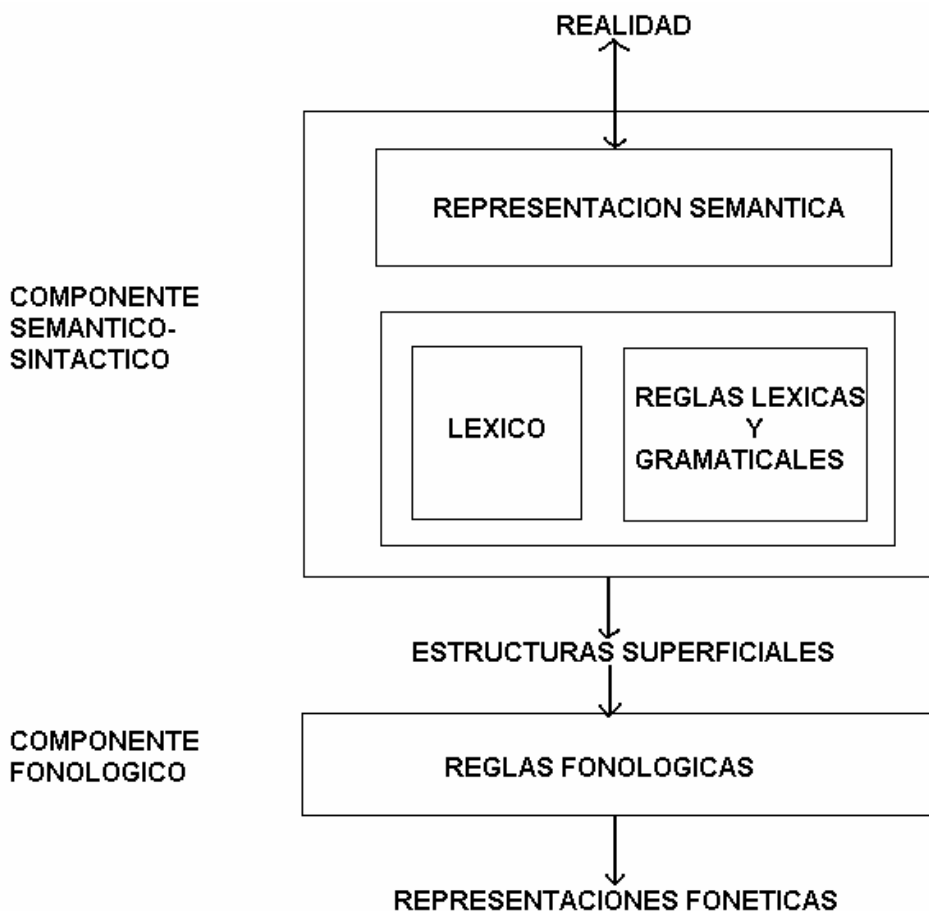
Respecto a los cuatro últimos componentes, aunque en principio tienen lugar de manera secuencial, una vez que se ha completado el primer ciclo de desarrollo, se da una permanente retroalimentación entre ellos, posibilitando así la cualificación de cada uno.

Descripción de los componentes investigativos.

1. Identificación y descripción de la perspectiva teórica que fundamenta la propuesta.

Componente decisivo para el desarrollo de la investigación, pues es el que da unidad y coherencia a toda la propuesta, ya que de la perspectiva teórica que se tenga acerca de un fenómeno, en este caso el del lenguaje, en general y el de la lengua, en particular, depende cualquier reflexión o desarrollo que se haga sobre tal objeto de estudio. Se trata entonces de ubicarse en una de las diferentes corrientes de pensamiento que han surgido durante la evolución de la lingüística como ciencia.

En este sentido, la perspectiva teórica que fundamenta la investigación desarrollada es la Semántica Generativa, modelo derivado del modelo estándar de Chomsky, por cuanto logra una explicación coherente y robusta acerca de los fenómenos lingüísticos implicados en la investigación, como lo es el proceso de generación de enunciados por parte del hablante de una lengua. Además, la Semántica Generativa ha posibilitado el desarrollo de modelos de análisis lingüístico que satisfacen las necesidades investigativas en términos de lograr una explicación cada vez más próxima y rigurosa frente a los diversos fenómenos relacionados con el lenguaje.



Segunda época, No. 11 – Segundo Semestre de 1999

Fig. 2 El Modelo de la Semántica Generativa

2. Selección y descripción del modelo de análisis lingüístico.

Se trata de seleccionar un modelo que represente el fenómeno lingüístico que se busca simular en el computador, a saber: la generación de enunciados simples por parte de un hablante. Este modelo debe estar en correspondencia con la perspectiva teórica e incorporar las variables que permitan explicar el fenómeno de una manera coherente y adecuada.

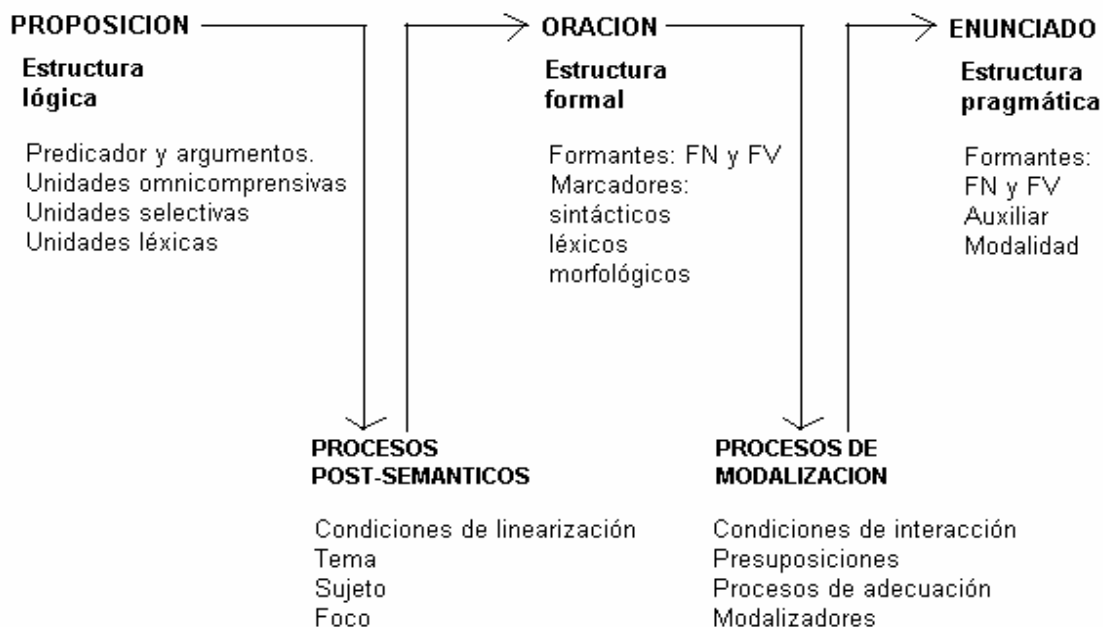


Fig.3. Modelo Semántico-pragmático.

Es así como se seleccionó el modelo Semántico –pragmático³, modelo que representa los procesos que se dan al interior de la mente, cuando un hablante genera enunciados verbales. Aunque el Modelo está inspirado en la Semántica Generativa, incorpora nuevos elementos tanto a nivel semántico como a nivel sintáctico y pragmático. Contempla, entonces, una unidad de análisis para cada uno de estos niveles. En el nivel semántico encontramos la proposición que es una unidad de carácter lógico o conceptual, en el nivel sintáctico tenemos la oración que es una unidad de carácter formal; por último tenemos el enunciado que es la unidad pragmática.

³ Este modelo fue propuesto y aplicado por la Dra. Lucía Tobón de Castro en el marco del curso de morfosintaxis española, en el Departamento de lenguas de la UPN. Aquí se trabajó con la versión vigente hasta el año 92, para posteriores desarrollos del modelo, el lector puede remitirse al tomo XLVIII de la revista *Thesurus* del año 1993.

3. Selección del lenguaje de programación

El lenguaje de programación seleccionado para desarrollar el software depende del grado de potencialidad que se desee, así como de las características que se espera que tenga tal desarrollo tecnológico. En el caso particular que aquí se presenta era necesario un lenguaje de programación que permitiera:

- a. *Desarrollar una programación modular*
- b. *Alcanzar un alto nivel de generalización*
- c. *Tomar decisiones a partir de inferencias*
- d. *Manipular de manera individual el mayor número de variables*

PROLOG, reúne todas estas características, además de que, al igual que el modelo de análisis lingüístico seleccionado, su estructura obedece al modo de funcionamiento de la lógica proposicional, lo que se constituye en una ventaja en el momento de traducir a reglas cada uno de los procesos implicados en la generación de un enunciado. Por estas razones el GESE se desarrolló con este lenguaje de programación.

4. Diseño del modelo computacional

Con el fin de adaptar el modelo lingüístico seleccionado como base para el desarrollo del generador, se hace necesario el diseño de un modelo computacional que, a la vez que corresponda y guarde coherencia con dicho modelo lingüístico, sea adecuado al lenguaje de programación seleccionado, y responda a los requerimientos propios de la informática para manipular los componentes y variables relevantes en el modelo.

La figura 4 representa el modelo computacional que subyace al GESE.

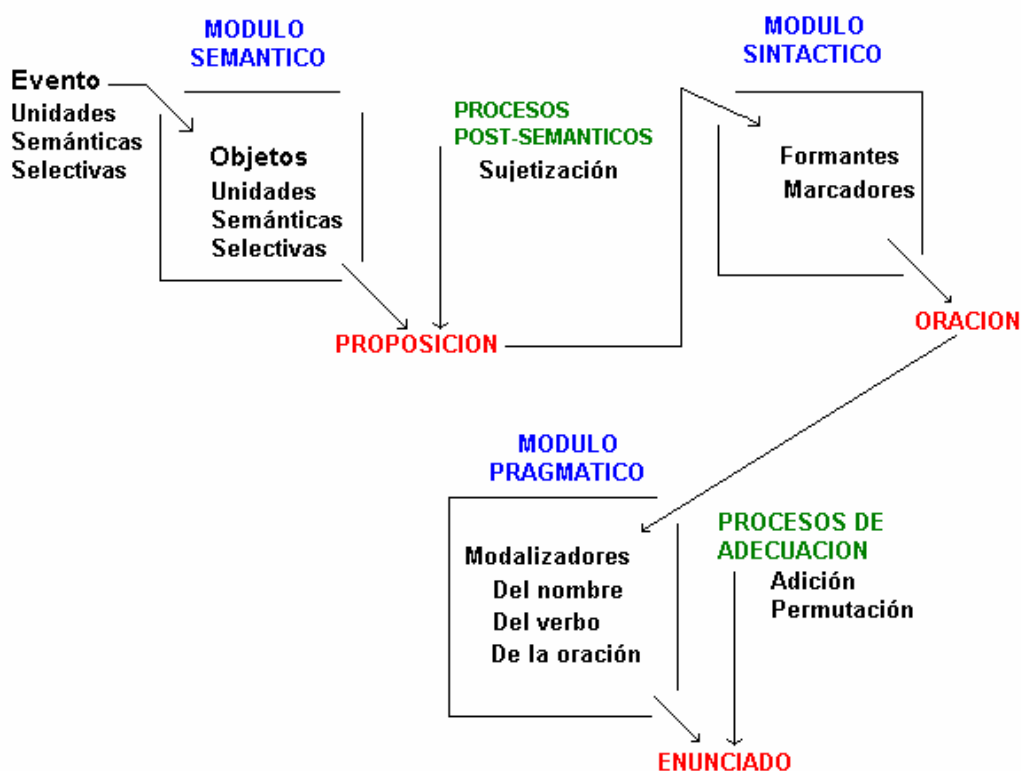


Fig.4. Estructura global del modelo computacional para la creación del GESE

5. Desarrollo del generador

Esta etapa del proceso investigativo corresponde al desarrollo tecnológico propiamente dicho; es el momento en el que confluyen todos los componentes anteriores, para concretizarse en un resultado tangible: el generador.

Se trata, entonces, de realizar la programación necesaria, en el lenguaje elegido y según los parámetros establecidos en el modelo computacional, de manera que se obtengan los resultados esperados. Es pues, el momento de mayor exigencia desde el punto de vista tecnológico, pues requiere de un buen manejo tanto del lenguaje como de estrategias y destrezas de programación, todo lo cual se debe conjugar, junto con el conocimiento del modelo lingüístico, para obtener un programa (software) efectivo, para dar origen al GESE como producto tecnológico concreto.

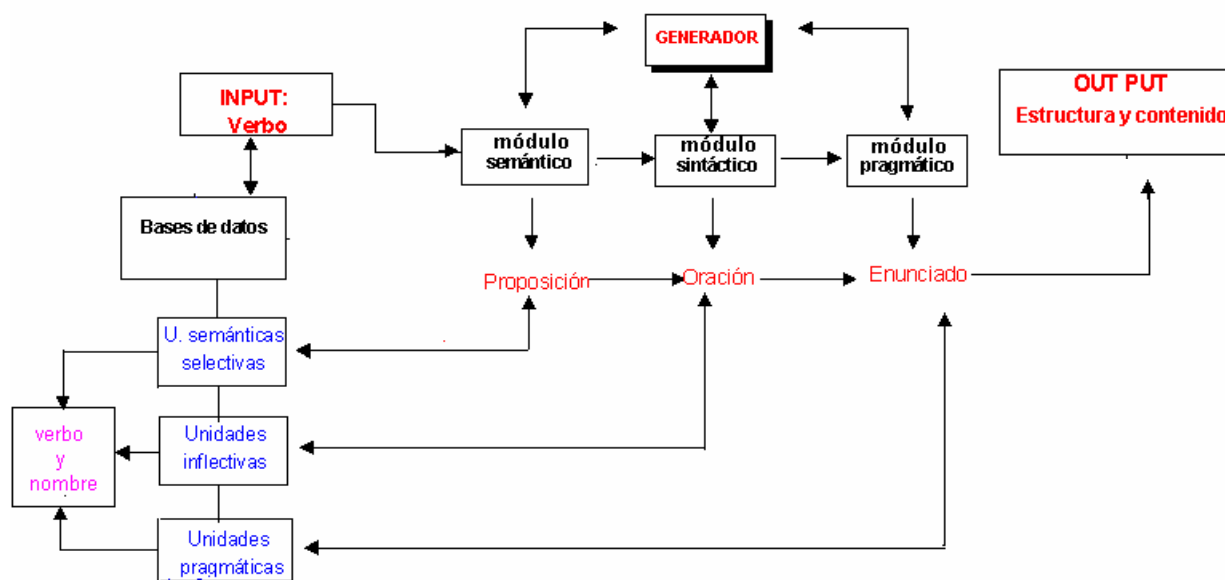


Fig. 5 . Estructura y funcionamiento del GESE.

Como puede verse, en este modelo, al igual que en el modelo lingüístico mencionado, está previsto que cada uno de estos módulos arroje como resultado una unidad lingüística: la proposición, la oración y el enunciado, respectivamente, a partir de un INPUT dado, conformado, en este caso por un verbo.

El GESE está desarrollado en PROLOG, versión 2.0, programa que facilita el procesamiento de lenguaje natural debido a su carácter de lenguaje de programación lógica, lo cual lo acerca a la lógica de funcionamiento de una lengua, en este caso, del español.

El GESE consta de una programación modular que facilita su manejo y agiliza su ejecución. Está construido como un "project" llamado, igualmente, gese. Este "project" está conformado por un total de 10 módulos, así:

Los módulos **uno**, **dos** y **tres** contienen toda la programación, por lo que constituyen el generador propiamente dicho. El módulo uno desarrolla los procesos correspondientes al componente semántico, el dos al sintáctico y el tres al pragmático.

Los módulos **baseuno**, **basedos**, **basetres**, **basecuatro**, **base4a**, **basecinco** y **base5a**: corresponden a bases de datos externas, para manejar la información correspondiente a los nombres y verbos. El programa inicia generando estas bases, por lo cual el goal se encuentra en el módulo **baseuno** que encadena los otros secuencialmente

Módulo **global**: contiene los dominios, predicados y bases de datos compartidos por todos los otros módulos.

Es obvio que el GESE, en su primer desarrollo, no puede generar un número infinito de enunciados simples, por una parte, y por otra, resultaría demasiado ambicioso pretender incluir absolutamente todas las variantes de enunciados que abarca el modelo. Por esta razón, se hizo necesario generar un espacio o micromundo, para delimitar su alcance.

A manera de ejemplo, y de forma sintética, veamos el resultado de cada uno de los módulos del GESE, teniendo como "input" el verbo *Viajar*.

Input: *viajar*

Output del módulo semántico:

❖ Estructura:

int(afir),prop([rol(verbo,viajar,[acc]),rol(agente,muchacha,[cont,anim,huma,fem]),rol(fin,cuba,[unico,espa,noprox])])

❖ Contenido léxico

viajar muchacha cuba

Output del módulo sintáctico

❖ Estructura:

(int(afir),oración([fn(nombre(muchacha,[fem,nopro,noplu,tp],[cont,anim,huma,fem]),fver([verbo(viajar,[infinitivo],[acc])),[fadv(fprep(prepare,fn(nombre(cuba,[fem,nopro,noplu,tp],[unico,espa,noprox])]))]))))

❖ Contenido léxico

muchacha viajar cuba

Output del módulo pragmático:

❖ Estructura:

int(afir),enunciado1([frase_nom(art(la,[def,espe,cono,fem,norpo,noplu,tp]),nombre(muchacha,[fem,nopro,noplu,tp],[cont,anim,huma,fem])),aux(tiempo(pretérito),aspecto(imperfecto,_),modal(obligativo,tuvo que)), frase_ver([ver(viajar,hecho anterior,final de acción,obligatoriedad],[noplu,tp]),[frase_adv(frae_prep(prepare),frase_nom(art(vacio,[vacio],nombre(cuba,[fem,nopro,noplu,tp],[unico,espa,noprox])]))))

❖ Contenido léxico

La muchacha tuvo que viajar a cuba

6. Validación.

La validación del modelo computacional y en consecuencia del generador, se realiza bajo dos parámetros: por una parte, el concepto de un experto, y por otra, su funcionamiento mismo, es decir, que en realidad ejecute sus tareas.

Estos dos criterios, son suficientes para avalar un desarrollo tecnológico de este tipo.

En el caso del GESE, tanto su coherencia con los planteamientos teóricos que lo fundamentan, como su concepción y funcionamiento tecnológico, fueron validados por dos expertos: el experto temático y el experto en programación, ambos directores del proyecto: El Dr. Luis Facundo Maldonado, como experto en programación y desarrollo tecnológico, y la Dra. Angela Camargo, experta en Lingüística, en general, y en el modelo Semántico-pragmático, en particular.

Por otra parte, el adecuado funcionamiento de la herramienta tecnológica fue verificado, utilizando el software con los diferentes "inputs", con lo cual se pudo constatar que el GESE genera enunciados simples en español, presentando tanto su contenido léxico, como su estructura en los tres componentes de análisis lingüístico: semántico, sintáctico y pragmático.

7. Conclusiones

Es en este componente del proceso investigativo en el que se explicitan los logros del mismo, tanto en términos del desarrollo tecnológico, como de los aportes que se hayan podido generar para las áreas de conocimiento comprometidas en el mismo: la lingüística aplicada y la lingüística computacional.

Ante todo, debe decirse que esta investigación, centrada en un desarrollo formal y tecnológico, cumplió tanto con sus objetivos tanto generales como con su objetivo específico. En efecto, se logró desarrollar un modelo computacional para desarrollar un software generador de enunciados simples en español a partir del modelo computacional basado en el modelo Semántico - pragmático de la lingüística aplicada, con las características requeridas: presentar tanto el contenido léxico como la estructura de los enunciados en los tres niveles: semántico, sintáctico y pragmático.

A continuación se presentan las conclusiones obtenidas a través del proceso investigativo, y que apuntan a dos áreas:

- **La lingüística computacional**
- **La lingüística aplicada**

7.1. Referidas a la lingüística computacional

Dado que desde esta perspectiva el trabajo estuvo centrado en dilucidar cuál era la vía más adecuada para crear el modelo computacional, la respuesta empezó a construirse desde el momento en que se eligió PROLOG como lenguaje de programación para desarrollar el GESE, pues esta elección implicaba que lo que se debería hacer era convertir en reglas cada uno de los pasos que, según dicho modelo, tienen lugar en el proceso de generación de un enunciado, como en efecto se hizo. Aquí el mayor trabajo radicó en buscar la forma de potencializar al máximo tales reglas, de manera que permitieran, por una parte la manipulación de todos y cada una de las variables que intervienen en la generación del enunciado y por otra, alcanzar el mayor grado de generalización.

Aquí debe decirse que los caminos probados fueron muchos, hasta encontrar que la mejor opción era trabajar cada una de las unidades de análisis (la proposición, la oración y el enunciado) como una lista de listas, recurso que permitió satisfacer los dos requisitos antes mencionados.

En cuanto al aporte del modelo computacional, y por ende del generador, a esta área de investigación, debe puntualizarse lo siguiente:

- a. El GESE es el único generador de enunciados simples en español basado en el modelo semántico pragmático y por lo tanto una auténtica alternativa para este tipo de desarrollos.
- b. Al usar dicho modelo como base, se garantiza la inclusión de variables que otros generadores no habían tenido en cuenta, como las matrices de rasgos semánticos y pragmáticos con el sentido que aquí se trabajan.
- c. Incorporación de la estructura de los diferentes niveles de análisis, además de la presentación de los contenidos del enunciado, lo que le da su carácter de recurso didáctico .
- d. Debido a la versatilidad de su programación, con solo modificar o ampliar sus bases de conocimiento, el GESE puede producir enunciados referidos a otros dominios.

7.2. Referidas al modelo Semántico- pragmático de la lingüística aplicada

Al traducir a reglas cada uno de los subprocesos propuestos por el modelo Semántico - pragmático, se evidenciaron algunos hechos sobre las siguientes características del mismo:

a. **La coherencia del modelo:** alcanza un alto nivel, lo que resulta muy valioso para desarrollar un programa de computador basado en la lógica, como lo es el GESE, hecho que puede deducirse fácilmente si se tienen en cuenta sus dos fuentes de desarrollo: el lenguaje natural y PROLOG.

b. **La completez:** el modelo permite manipular un gran número de las variables que intervienen en el momento de generar un enunciado. Sin embargo, al desarrollar el programa computacional surgió la necesidad de hacer explícitos algunos hechos que en el MSP se manejan de manera implícita, posiblemente porque este modelo está pensado como instrumento didáctico, lo que supone que es trabajado por un hablante real que sabe de antemano cuándo un enunciado es inadecuado; cosa que no ocurre en el modelo computacional, pues el programa debe trabajar sólo y por lo tanto debe alimentarse de la forma más detallada posible para evitar enunciados inadecuados.

Se identificaron, en este sentido, los siguientes casos:

a). El programa debe tener información explícita y suficiente para decidir cuándo es adecuada la relación entre un verbo determinado y un modal, pues no todos los verbos aceptan todos los modales.

b). El programa debe poseer información explícita acerca de la relación que hay entre el formante sintáctico y asignado a un rol semántico y el artículo modificador del nombre que cumple ese rol. Así, por ejemplo, un locativo temporal formado por el nombre +femenino y con un formante sintáctico del tipo:

FAdv ---> FP
 FP ---> FN
 FN ---> Art N

solamente aceptará como modificador del nombre (N) un artículo definido-definido (la), para obtener construcciones como:

Compré la leche por **la** tarde

De no hacerse explícita esta restricción, se generarían enunciados anómalos como:

Compré la leche por **esta** tarde.

c) Dentro del MSP se trabaja con la Unidades semánticas Selectivas del nombre y del verbo y con las Unidades Inflectivas del nombre. En el modelo computacional,

es, en el GESE surgió la necesidad de explicitar además las unidades inflectivas del verbo, conformadas por información referida al número (+ ó - plural) y a la persona (primera, segunda, tercera). Estas unidades son definitivas para establecer la concordancia verbal.

Además, en el GESE se introduce el concepto de Unidades Pragmáticas Selectivas (UPS) tanto del nombre como del verbo, para representar los rasgos que el hablante, según su intencionalidad y el contexto del momento de la enunciación, tiene en cuenta para decidir, por una parte, el artículo modalizador del nombre, y por otra, los componentes del auxiliar como modalizador del verbo.

La matriz de las UPS del nombre está conformada por dos de los rasgos que en el MSP se caracterizan como Unidades Inflectivas: conocido y específico, más el rasgo que proporciona información acerca de la relación que hay entre el referente y el hablante en términos de distancia temporo-espacial y de pertenencia.

En cuanto a las UPS del verbo son rasgos que proporcionan información referida a la relación entre el momento de habla y el momento de la acción, a la referencia al inicio, desarrollo o final de la misma, y a la presencia o ausencia de un matiz semántico modificador del verbo principal.

d) En el modelo computacional se hace más explícita la fusión de elementos de los tres niveles de análisis (semántico, sintáctico y pragmático) en el momento de generar un enunciado, pues hay información del nivel semántico, por ejemplo, que debe ser almacenada hasta el nivel pragmático para evitar enunciados anómalos, tal es el caso de la Unidad Semántica Selectiva +femenino, que debe acompañar al nombre hasta el último nivel para establecer la concordancia nominal y verbal.

c. **La secuencialidad de los subprocesos:** El algoritmo de generación de enunciados propuesto por el MSP se modificó en lo referido al proceso post-semántico de tematización, a los procesos de adición y permutación y a la modalización de ambiente de los verbos de acción, de la siguiente manera:

Proceso post-semántico de tematización: el MSP lo trabaja entre el nivel semántico y el nivel sintáctico. En el GESE se trabaja en el nivel pragmático, junto con el proceso de permutación, ya que está relacionado con él.

Procesos de adición y de permutación: el MSP los trabaja en el análisis sintáctico. En el GESE se abordan en el nivel pragmático.

Proceso de modalización del ambiente en verbos de acción: el MSP lo trabaja en el nivel pragmático. En el GESE se incorpora desde el nivel semántico.

La razón para estos cambios está dada en términos de ganancia en economía y agilidad en la programación, ya que resultaba más práctico trabajar primero la generación de un enunciado canónico en español para, a partir de él, derivar

enunciados atípicos. De la misma manera, resultaba más apropiado trabajar primero estructuras generales para luego producir las particulares.

Si se tiene en cuenta que estos son cambios de tipo formal es posible entender que no implican contradicciones conceptuales entre el MSP y el GESE. Lo que puede evidenciarse aún más si se tiene presente que la secuencialidad de estos procesos es un elemento de carácter didáctico y científico, ya que en un acto real de habla todos ellos se dan simultáneamente.

La gran conclusión a que nos lleva este análisis y que hace parte de la respuesta al interrogante inicial, es que el Modelo Semántico - Pragmático sí resulta ser un instrumento adecuado, potente y versátil para el procesamiento del lenguaje natural, al menos en lo concerniente con el proceso de generación, como lo demuestra el hecho de haber desarrollado un generador de enunciados simples basado en dicho modelo.

De esta forma el GESE como producto de un proceso de investigación y de un desarrollo tecnológico, contribuye en principio, al estudio de, por lo menos, tres áreas del saber:

1. La lingüística aplicada: porque plantea un nuevo campo de aplicación de un modelo lingüístico, el semántico-pragmático.
2. La lingüística computacional: porque pone al servicio de ella un modelo lingüístico, el semántico pragmático, como una nueva alternativa para el procesamiento de lenguaje natural. Además, porque es un producto nuevo, ya que la mayoría de sus homólogos son generadores de enunciados o de textos en lengua inglesa. Su novedad también radica en que no se limita a producir los enunciados, sino que hace explícitas y manipulables sus estructuras semántica, sintáctica y pragmática.
3. La didáctica de la morfosintaxis del español: porque el programa permite analizar y comparar las estructuras de diferentes enunciados, en los tres niveles de análisis: semántico, sintáctico y pragmático.
Por último, debido a la versatilidad de su programación, con solo modificar o ampliar sus bases de conocimiento, el GESE puede producir enunciados referidos a otros dominios.

De esta manera, conduye este trabajo con la certeza de haber plasmado en él lo mejor de mi formación profesional y humana, para contribuir de alguna manera al desarrollo del país.

8. Proyecciones

La investigación aquí presentada tuvo su origen en el marco de un proyecto más global, que respondía a la esencia de la Maestría en cuestión : aplicar el uso de tecnologías de la información a procesos pedagógicos y/o de aprendizaje. Tal proyecto tenía como objetivo desarrollar un tutor inteligente para el aprendizaje de la morfosintaxis del español, para lo cual era indispensable crear, en primera instancia, un generador inteligente que a la vez que utilizara, como era obvio, la morfosintaxis del español para generar sus enunciados, la hiciera explícita y posibilitara la manipulación de los diferentes elementos que la componen. Es en respuesta a este requerimiento como nace el GESE y el proceso investigativo que lo soporta.

Teniendo en cuenta lo anterior, la principal proyección de esta investigación, consiste en robustecer el GESE hasta convertirlo en un tutor inteligente para el aprendizaje de la morfosintaxis del español, proceso indispensable en las facultades de lingüística y, en algunas ocasiones, para la apropiación del español como lengua extranjera por parte de adultos.

Sin embargo, aunque en esta primera versión el GESE no tiene las características de un tutor inteligente, ya posee herramientas que permiten utilizarlo en procesos de aprendizaje de la morfosintaxis del español, pues, como se ha enfatizado, el GESE no se limita a producir los enunciados, sino que muestra la estructura de los mismos en sus tres niveles de análisis (semántico, sintáctico y pragmático)!. Este solo elemento constituye ya un valioso recurso didáctico, por cuanto permite, por ejemplo, hacer análisis comparativos a partir de los cuales un aprendiz puede deducir ciertas reglas de la morfosintaxis del español.

Por último, es conveniente adarar que, si bien la proyección antes descrita es la que responde a la intencionalidad con la que fue desarrollado en GESE, no es la única posible, pues el GESE en una segunda etapa de investigación también podría convertirse en una herramienta para generar las formas verbales del planteamiento de problemas, en una interface en lenguaje natural de otro programa, en fin, en software complementario de cualquier programa que requiera de la generación de enunciados simples en español.

BIBLIOGRAFIA

- BATEMAN, MOORE y WHITNEY. **Upper modeling: A level of semantics for natural language processing**. En IWNLG, editor, Proceedings of the Fifth International Workshop on Natural Language Generation, Pittsburgh, Pennsylvania, Abril 1990. Springer-Verlag.
- BODDARD, Helen. **Cognitive Modelling of Language acquisition and generation**. University of Exeter, 1985.
- CONTRERAS, Heles. **El Orden de las Palabras en Español**, siglo XXI, 1978
- CHAFE, Wallace. **Significado y Estructura de la Lengua**. Planeta, 1976.
- CHOMSKY, Noam. **Syntactic Structure**. Mouton, 1957
- DALE, HOVY, RÖSNER y STOCK, editores. **Aspects of Automated Natural Language Generation**. Number 587, en Lecture Notes en AI. Springer-Verlag, Heidelberg, 1992.
- DALE, MELLISH y ZOCK, editores. **Current Research in Natural Language Generation**. Academic Press, London, 1990.
- FILLMORE, Charles. **Hacia una Teoría de Casos**. En Los Fundamentos de la Gramática Transformacional de Helles Contreras, Siglo XXI, 1974.
- HOVY, Eduard. **Language Generation**. University of Southern California, Marina del Rey, California, 1997.
- JACOBS, P. S. **Natural Language Generation: Recent Advances in Artificial Intelligence, Psychology, and Linguistics**. En COLING [COL96] Gerard Kempen, editor. Kluwer Academic, Boston, Dordrecht, 1996.
- KATZ, Jerrald y otros. **La Estructura de la Teoría Semántica**. En Los Fundamentos de la Gramática Transformacional de Helles Contreras, Siglo XXI, 1974.
- KOWALSKY, Robert. **Lógica, Programación e Inteligencia Artificial**, Ed. Díaz de Santos, Madrid, 1986.
- LEON ORTIZ, M. IDALITH. **Modelos lingüísticos**. Universidad del Tolima, Ibagué, 1995.
- LLOPART et al. **Lingüística Computacional**, 1990.
- MALDONADO, Luis Facundo. **Programación Lógica**. Bogotá, 1993.
- NEUMANN, Günter. **A Uniform Computational Model for Natural Language Parsing and Generation**. PhD thesis, Universität des Saarlandes, Germany, Noviembre 1994.
- RODRÍGUEZ-ROSELLÓ, **Introducción a los micromundos matemáticos**, 1986.
- SHAPIRO, E. **Language Understanding and Generation**. En Encyclopedia of Artificial Intelligence, pag 630-, 1987.
- SCHANK, R.C. **Conceptual information processing**. North-Holland, New York, 1975.
- SOWA, J.F. **Conceptual structures: information processing in mind and machine**. Addison – Wesley, Reading, 1984.
- TOBON DE CASTRO, LUCIA. Documentos producidos para la cátedra de Morfosintaxis del español en la Universidad Pedagógica Nacional. Bogotá, 1985-1989.
- ___ **Propuesta de un proyecto lingüístico para el estudio del español como lengua propia**. En revista Thesaurus, tomo XLVIII, Instituto Caro y Cuervo, Santafé de Bogotá, 1993.

TURBO PROLOG 2.0 REFERENCE GUIDE.

VAN DIJK, TEUN. **Estructuras y funciones del discurso**. México, Siglo XXI, 1989.

VANNOORD ,G. **An overview of head-driven bottom-up generation**. En Robert Dale, Chris S. Mellish, and Michael Zock, editors, Current Research in Natural Language Generation, pages 141--165. Academic Press, London, 1990.

DIRECCIONES ELECTRÓNICAS

<http://www.kgw.tu-berlin.de/~mengel/SpeechTech/ch4node2.html>

<http://www.dsv.su.se/~hercules>

<http://www.itri.brighton.ac.uk/posts/gnome-advert.html>

<http://iinwww.ira.uka.de/bibliography/Ai/nlg.html>

http://www.dfki.unisb.de/fluids/Language_Generation_Research_and_Market_Trends_in_Japan_and.html

<http://www.syd.dit.csiro.au/staff/cecile/esslli/schedule.html>

<http://www.hcrc.ed.ac.uk/Site/RTNLGREA.html>

